REMARKS

Claims 1, 4-9, 11-12, 14, 17-22, 24-27, 29, 31-34, 36-37, 40-44, 47-55, and 57 are pending. All those claims are rejected as being anticipated over the reference of Burnett, et al., U.S. Patent Application Publication No. 2003/0228023. Claims 7, 9, 12, 18, 20, 25, 27, 29, 37, and 49 are also rejected under 35 U.S.C. §112 as being indefinite.

Rejections Under 35 U.S.C. §112

The claims noted above, that have been rejected with respect to 35 U.S.C. §112 have all been amended to recite "the baseline", as opposed to "a baseline". The baseline referred to is set forth in a previous independent or dependent claim.

Rejections Under 35 U.S.C. §102

All the pending claims are rejected over the reference of Burnett, et al. `023. Also, the Examiner makes references to a companion case, including Burnett, et al., U.S. Patent Application Publication No. 2003/0179888, and thus, both references are addressed herein.

The invention recited in the claims, as currently amended, is directed to an apparatus for use in capturing and detecting speech for the purposes of speech recognition processing. As such, the invention involves selectively performing speech recognition, or selectively not

18

performing speech recognition, based upon the operation and various features of the invention. The present invention utilizes a plurality of microphones, such as first and second microphones, to capture different audio signals. The processing circuitry of the invention then performs an analysis of the captured audio signals to determine if the user is speaking. The processing circuitry is configured for selectively forwarding audio signals from one of the microphones, such as the first microphone, to speech recognition circuitry, for further speech recognition processing. However, the processing circuitry only forwards those signals from the first microphone, when the analysis indicates that the user is speaking. The processing circuitry does not forward the audio signals to the speech recognition circuitry, and further speech recognition processing is not completed when user speech is not detected. The current independent claims pending in the Application have been amended to further clarify the issue, and to address the Examiner's comments on arguments set forth in the last Response to Office Action in this Application. Furthermore, the claims are amended to particularly clarify the features of the invention directed to the selective speech recognition processing that the invention provides, which is not taught or suggested by either of the Burnett, et al. `023 or `888 references.

Certainly, those Burnett, et al. references show systems that utilize multiple microphones, such as first and second microphones. Furthermore, those microphones may be utilized to operate in conjunction with a voice activity detector (VAD) to determine if there is speech or no speech. However, any similarities end there. In the Examiner's response to the earlier submitted arguments, the terms "selectively forwarded" or not "forwarded" are set forth with quotation marks as if to assert that such limitations, as specifically recited in the pending claims, are somehow to have no effect with respect to an analysis of the claims under 35 U.S.C. §§ 102 and 103 of the Patent Statute. However, such limitations must be considered. It is the Applicant's hope that the further Amendments to the claims to clarify the operation and function of the invention will further help distinguish over the cited art.

The Examiner argues that it is well known that the primary function of a voice activity detector (VAD) is to provide an indication of speech presence in order to facilitate additional and/or further speech processing. The Examiner then further argues that a signal, which is determined to be voice/speech is "selectively forwarded" for further speech processing, and a signal which is determined to be non speech is not "forwarded", and additional speech processing is not initiated. However, while it is true that a voice activity detector (VAD) provides an

indication of the presence of speech, it is not used to selectively control speech recognition processing, such that further speech recognition processing is addressed when the user is speaking, but is not completed when user's speech is not detected, as recited in the currently-pending Claim 1, for example.

The Examiner makes reference to "further speech processing" as if to somehow assert that anything that is done after using the voice activity director (VAD) is somehow selectively controlled by that VAD operation. That is simply untrue and, in fact, is technically wrong.

The Burnett, et al. `023 reference specifically sets forth the purpose of the VAD. Referring to Paragraph [0039] in the Burnett, et al. `023 reference, it clearly states that the VAD provides "voice activity signals that include information of human voicing activity". Then, "Components of the communication systems use the control signals (of the VAD) to automatically select a de-noising method appropriate to . . . the acoustic signals" [Paragraph 0039]. That is, the VAD signal is used to select the type of de-noising that might be utilized to clean up the audio signal. As set forth in the Response to the earlier Office Action, the Burnett, et al. `023 reference is directed to de-noising, and is particularly directed to the types of microphones that might be used for de-noising, and their respective positions.

Furthermore, in Paragraph [0135], the particular purpose of the VAD is further discussed with respect to the prior art of classical adaptive noise cancellation (ANC). Specifically, as set forth in Paragraphs [0135] and [0136], the VAD information is used to control adaptation of the noise suppression system. More specifically, in making reference to the Burnett, et al. `888 reference, which further discussed the VAD and methods for use, the VAD is used to modify the transfer functions $H_1(z)$ and $H_2(z)$ of the de-noising system. Paragraph [0136] of the Burnett, et al. `023 reference cites "therefore, the various methods described above use VAD information to construct a sufficiently accurate VAD to instruct the Pathfinder system when to adapt the coefficients of H_1 (noise only) and H_2 (if needed when speech is being produced).

The Burnett, et al. `888 reference, which goes into further detail with respect to the use of two microphones and the generation of a VAD signal within a larger Pathfinder noise suppression system, specifically states in Paragraph [0041] and Paragraph [0044] that the Pathfinder system uses the VAD information and the received acoustic information from the multiple microphones to reduce or eliminate noise in desired acoustic signals by estimating the noise wave form, and subtracting it from a signal including both speech and noise. With respect to how

speech or the absence of the presence of speech affects that system,

Paragraph [0044] further states:

> When the VAD signal 104 indicates an
> absence of voicing, the Pathfinder system
> 101 uses MIC 1 and MIC 2 signals to
> calculate the coefficients for a model of
> transfer function $H_1(z)$ over pre-specified
> subbands of the received signals. When the
> VAD signal 104 indicates the presence of
> voicing, the Pathfinder system 101 stops
> updating $H_1(z)$ and starts calculating the
> coefficients for transfer function $H_2(z)$ over
> pre-specified subbands of the received
> signals. (Paragraph [0044]).

That Paragraph [0044] also goes on to state how the Pathfinder

system then utilizes the $H_1(z)$ and $H_2(z)$ transfer functions to generate a

de-noised acoustic stream.

Therefore, the primary function of the VAD, as discussed in both

the Burnett, et al. `023 and `888 references is to affect the method of de-

noising used in the overall noise suppression system, and also to update

the coefficients of the transfer functions that are used for de-noising.

There is absolutely no teaching or discussion within either of those

references regarding processing circuitry that is configured for

selectively forwarding audio signals from a first of the microphones to

speech recognition circuitry for further speech recognition processing

only when the audio signals indicate that the user is speaking, but not

forwarding the audio signals from the first microphone to the speech

recognition circuitry, and not completing the further speech recognition

processing when the user's speech is not detected.

Burnett, et al. `023 and `888 do not, in any way, affect how the speech recognition processing is selectively used or not used. Rather, those references are directed to noise suppression from an acoustic stream or audio stream. The functionality of the present invention, with respect to the selective control of further speech recognition processing is a specific limitation, as set forth in the independent claims of the Application, and particularly independent Claim 1. Such a limitation cannot simply be dismissed as if it has no meaning. Furthermore, it cannot be generically dismissed by mischaracterizing it. In the Burnett, et al. references, the cleaned-up speech signal always proceeds to whatever its final destination is. The use of the VAD may determine how much noise is actually removed, but the signal proceeds regardless. It is never stopped or prevented from further processing. Nor is it selectively subjected to speech recognition processing. Therefore, the Examiner's assertion that the speech signal is selectively forwarded or not forwarded is simply in error.

Turning now to the specific rejections, the Examiner cites to an incredibly large number of paragraphs within the Burnett, et al. `023 reference to somehow provide the teaching of the recited claim elements in the pending case. Specifically, the Paragraphs [0043], [0051-0052], [0087-0136], and [0142-0162] are all recited for teaching the processing circuitry elements set forth in pending Claim 1. However, nowhere in any of those numerous paragraphs is there a teaching of the limitations

set forth with respect to the processing circuitry in Claim 1. Rather, most

of those paragraphs discuss different VAD methodologies. Again,

Applicant is not arguing that the Burnett, et al. reference does not utilize

a VAD that indicates the presence of speech or no speech. Rather,

Applicant submits, as set forth in detail above, that there is no teaching

in the cited art regarding processing circuitry that is configured for

selectively forwarding audio signals from one microphone for further

speech recognition processing, when it is determined that the user is

speaking, but not forwarding the audio signals from the first microphone

for further speech recognition processing by the speech recognition

circuitry when user speech is not detected. The various cited

paragraphs generally teach the following:

Paragraph [0043] – This paragraph does not at all discuss selectively
forwarding audio signals to speech recognition circuitry. The paragraph
only discusses the VAD, or voice activity detector, which indicates that
speech has occurred or not occurred during a particular time sample of
audio.

Paragraph [0051] – This paragraph only discusses that the VAD may be
particular sensors, such as accelerometers, skin surface microphones,
etc. It also references U.S. Application No. 10/383,162.

Paragraph [0052] – This paragraph states that configurations, as
described in the application, are directed to the location and orientation of
the microphones, and the method used to maintain a reliable VAD signal.
There is no discussion regarding selective forwarding of audio signals.

Paragraphs [0087] and [0088] – These paragraphs merely generally
discuss VAD devices, and note the patent application, Serial #10/383,162.

Paragraphs [0089-0093] – These paragraphs merely discuss a general electro-magnetic sensor VAD. Such a VAD is used to detect vibrations of the head and neck associated with the production of speech. It detects areas of sensitivity to detect the vibrational signals associated with voicing. The VAD is an RF sensor that uses antennas for positioning of the sensors on the human head and neck, as discussed. There is no discussion whatsoever with respect to selective forwarding or speech recognition of audio signals.

Paragraphs [0094-0097] – These paragraphs discuss the surface skin vibration VAD. They detect skin vibrations that occur due to the production of speech. Accelerometers might also be utilized. A VAD does not require high fidelity, as only the ability to determine if vibrations are taking place is important. During speech when the accelerometer/SSM is placed on the cheek or neck, vibrations associated with speech are detected. The acoustic signal is used to generate a VAD signal that is then used to process and de-noise the signal of interest.

Paragraphs [0098-0099] – These paragraphs are directed to skin vibrations in the ear for generating the VAD signal. No discussion of selective processing.

Paragraphs [0100-0103] – These paragraphs are directed to skin vibrations outside of the ear.

Paragraphs [0104-0116] – These paragraphs describe other VAD's for use in the invention. For example, in a stereo VAD, the difference in frequency amplitude from the noise and the speech is used to determine when speech is occurring.

Paragraphs [0117-0131] – These paragraphs discuss how to determine the transfer functions within the system that are used for generating the VAD signal. The VAD signal is derived in some manner, and is used to control the method of noise removal in the overall noise suppression system.

Paragraphs [0132-0136] – These paragraphs discuss the use of VAD to control adaptation of the noise suppression system. The various methods used in the VAD information to construct a sufficiently accurate VAD to instruct the Pathfinder system when to adapt the coefficients of the transfer functions.

Paragraphs [0142-0146] – These paragraphs discuss the use of VAD to calculate the transfer functions and also discuss that the VAD may have errors that may affect the de-noising algorithm. There is no discussion of selectively forwarding the audio signals from the first microphone to the

speech recognition circuitry as claimed. The purpose of the VAD signal is to affect the de-noising algorithm. It does not selectively forward a signal for further speech processing, or selectively not forward such a signal.

Because the Burnett, et al. `023 reference does not teach each and every one of the limitations, as set forth in Claim 1, that claim is not anticipated under Section 102 and is allowable over the cited art. Dependent Claims 4-9 and 11-12 each depend from Claim 1, and further recite unique combinations of elements, which are not taught by the cited art. Accordingly, those claims are also allowable.

Independent Claim 14 and dependent Claims 17-21; independent Claim 22 and its dependent Claims 24-27 and 29; independent Claim 31 and its dependent Claims 32-34, 36-37, and 40-43; independent Claim 44 and its dependent Claims 47-51 and independent Claim 52 and its dependent Claims 53-55, and 57 are also in allowable form for the same reasons as discussed above, with respect to independent Claim 1 and its various dependent claims.

In light of the foregoing, it is respectfully submitted that the present application is in a condition for allowance and notice to that effect is hereby requested. If it is found that the present amendment does not place the application in a condition for allowance, Applicant's undersigned attorney requests that the examiner initiate a telephone interview to expedite prosecution of the application.

Applicant is submitting the fee due for the one-month extension of time with this response. If any additional fees are necessary, the Commissioner may consider this to be a request for such and charge any necessary fees to deposit account 23-3000.

Respectfully submitted,

WOOD, HERRON & EVANS, L.L.P.

/Kurt A. Summe/
Kurt A. Summe
Reg. No. 36,023

2700 Carew Tower
441 Vine Street
Cincinnati, Ohio 45202
(513) 241-2324
(513) 241-6234 facsimile

Document #477859